

# **Item response theory analysis applied to the Spanish version of the Personal Outcomes Scale**

J. Guàrdia-Olmos,<sup>1</sup> M. Carbó-Carreté,<sup>2</sup> M. Però-Cebollero<sup>1</sup> & C. Giné<sup>2</sup>  
1 Faculty of Psychology, Institute for Brain, Cognition and Behaviour (IR3C),

University of Barcelona, Barcelona, Spain

2 Faculty of Psychology, Education and Sport Sciences, Blanquerna, Ramon Llull  
University, Barcelona, Spain

## **Abstract**

### **Background**

The study of measurements of quality of life (QoL) is one of the great challenges of modern psychology and psychometric approaches. This issue has greater importance when examining QoL in populations that were historically treated on the basis of their deficiency, and recently the focus has shifted to what each person values and desires in their life, as in cases of people with intellectual disability (ID). Many studies of QoL scales applied in this area have attempted to improve the validity and reliability of their components by incorporating various sources of information to achieve consistency in the data obtained. The adaptation of the Personal Outcomes Scale (POS) in Spanish has shown excellent psychometric attributes, and its administration has three sources of information: self-assessment, practitioner and family. The study of possible congruence or incongruence of observed distributions of each item between sources is therefore essential to ensure a correct interpretation of the measure. The aim of this paper was to analyse the observed distribution of items and dimensions from the three Spanish POS information sources cited above, using Item Response Theory (IRT).

### **Method**

We studied a sample of 529 people with ID and their respective practitioners and family member, and in each case, we analysed items and factors using Samejima's model of polytomous ordinal scales.

### **Results**

The results indicated an important number of items with differential effects regarding sources, and in some cases, they indicated significant differences in the distribution of items, factors and sources of information.

## Conclusions

As a result of this analysis, we must affirm that the administration of the POS, considering three sources of information, was adequate overall, but a correct interpretation of the results requires that it obtain much more information to consider, as well as some specific items in specific dimensions. The overall ratings, if these comments are considered, could result in bias.

**Key Words:** Intellectual Disability, Quality of Life, Personal Outcomes Scale, Item Response Theory, Samejima's model

## Introduction

Personal outcomes related to quality of life (QoL) provide a framework to assess the impact of processes and services delivered to people with intellectual disabilities; moreover, they provide information about how people experience a positive and satisfying life (Reinders & Schalock 2014; Schalock *et al.* 2007; Schalock *et al.* 2008a,b)

Personal outcomes are "the aspirations defined and valued by the individual. This term is often used in general in relation to the dimensions and indicators of quality of life" (Schalock *et al.* 2008b, p. 278). The assessment of personal outcomes indicates consideration of the four principles underlying the measurement of the QoL construct: (1) assessing the degree to which people have life experiences that they value; (2) identifying the dimensions that contribute to a full and interconnected life; (3) involving the physical, social and cultural contexts that are important to people; and (4) including measurements of experiences that are common to all and of those that are unique to the person (Schalock & Verdugo 2002; Schalock *et al.* 2007). Based on these principles, the assessment reflects the subjective viewpoint (satisfaction and personal preferences) and objective information (experiences and personal circumstances).

This comprehension of the measurement of QoL has brought about the challenge of using self-report measures (Bonham *et al.* 2004; Li *et al.* 2013; Verdugo *et al.* 2005) given that incorporating the subjective perceptions is the way to grant an active role of people with ID in order to talk about their own lives (Cummins 2005; Schalock *et al.* 2002; Stancliffe, 2000). Nevertheless, the use of self-reports can involve some concerns related to problems in the comprehension of the items and for a high level of

acquiescence (Finlay & Lyons 2001, 2002; Hartley & MacLean 2006), and as consequence, it is more difficult to obtain reliable evaluations (Brown *et al.* 2013). Accordingly, a number of studies have tried to solve this issue and included the evaluation of a practitioner and/or a family member (Balboni *et al.* 2013; Claes *et al.* 2012a; Janssen *et al.* 2005; Perry & Felce, 2002, 2005; Schmidt *et al.* 2010; Schwartz & Rabinovitz, 2003; Simões & Santos 2016) These studies have emphasised the measurement of QoL from people with ID, as well as people who know about their functioning in different daily environments (practitioner and/or family).

The discrepancies observed between these perspectives have become a key point in the assessment of personal outcomes that can be used at an individual and organizational level (Schalock *et al.* 2008a,b, Reinders H.S. & Schalock 2014). At the individual level, the personal outcomes provide information about the experiences and circumstances impacting in the person's life, evaluated through the individual supports plan (ISP). The ISP serves as a guide to design and provide the individualised supports and evaluate if the desired goals and experiences are being accomplished based on enhanced personal outcomes (Thompson *et al.* 2009; van Loon, 2015). The framework of the ISP is the ecological model of disability and the different contexts where people's lives must be taken into account for a whole comprehension of their current QoL (Claes *et al.* 2012b, Reinders & Schalock 2014; Schalock *et al.* 2008a; van Loon, 2015). At the organizational level, personal outcomes could be useful for reporting whether the supports are meeting the needs of the people assisted, for evaluation and monitoring progress and improving quality services (Thompson *et al.* 2009; Schalock & Verdugo, 2012).

Based on this, the assessment of QoL requires that, in addition to the person with ID, the proxies also provide information on the environments where he or she is functioning. This approach has been reflected in the latest QoL scales (van Loon *et al.* 2008; Verdugo *et al.* 2013). Schalock & Felce (2004) argued that a correct approach to assessing QoL is based on a combination of objective and subjective measurements on the same scale. Later works emphasised that measuring subjective (self-report) and objective (report-of-others) factors using the same indicators and the same items resolved many of the problems associated with measurements focusing only on the objective or subjective part (Schalock *et al.* 2008a). Following this suggestion, the Personal Outcomes Scale (POS) (van Loon *et al.* 2008) become an adequate instrument to assess the QoL of people with ID as well as the Spanish version (Carbó-Carreté *et al.*

2015) that is used in the present work. Previous studies have been used the POS to evaluate the relationship between person with ID and their proxies (Balboni *et al.* 2013; Claes *et al.* 2012a; Simões & Santos 2016) and provide data referring the discrepancies and similarities in the dimensions scored.

The key lies in determining the similarities and discrepancies mentioned among the sources of information for different plausible scenarios showing the different contexts where the person functions. This matter may be approached in different ways. In some cases, the linear correlation between the assessments of different values was used, assuming that each source is assessing the same item. In the current paper, this is a disputable assumption, given that this is a self-assessment for the person with ID, whereas it is a hetero-assessment for the other two sources. Another possibility to approach this matter is comparing the values of means and variances between the same dimensions among the different sources of information. This would allow us to assess whether the different informants are presenting similar mean values regarding the intensity of the assessment. By no means will we be reporting on whether each source's response pattern presents discrepancies or congruences beyond the indicator of the mean. Finally, the alternatives to the Classic Test Theory (CTT), such as the Item Response Theory (IRT), allow us to overcome both limitations as we can analyse the distribution pattern of all the sources simultaneously, including the determination of the range of values each source presents for each item and the discrepancies and congruences to be expected in every case. Determining these patterns is essential to interpret the POS according to each informant. The congruences or non-congruences may be reasonable, but only as long as they fit into a known pattern which has been previously described. Establishing this pattern is the goal we pursue in the current paper.

Regarding the Spanish version of the POS, this study has not been undertaken, and obviously, the issue must be addressed. The IRT identifies curves characteristic of the items, which allow us to assess whether the three sources used show similarities and dissimilarities between the observed distribution between sources of the QoL measurement. Thus, the aim of this study was to estimate the basic IRT models appropriated to a categorical variable (sources used: self-reporting, reporting of practitioners and reporting of families) to assess the functioning among the aforementioned sources in studying the logistic distribution of the items.

## Method

### Sample

The sample consisted of a total of 529 people with ID (296 men and 233 women), with  $M_{age}=35.03$ ,  $SD=10.82$  (range 16-66 years old), who came from seven autonomous communities in Spain: Andalusia (20.9%), Aragon (4%), Catalonia (25%), Castile and León (6.6%), Castile-La Mancha (14.8%), Madrid (17.4%), and Galicia (11.7%). In addition, practitioners ( $n=522$ ) and the participants' families ( $n=462$ ) participated. In relation to participants' families the 71.9% was their father or mother, 21.6% was brother or sister; 3.2% was legal tutor and the rest of 3.2% was others. In relation to the gender, in all cases, 30% was male and 70% were female.

In this study, accidental, non-randomised sampling was performed in every autonomous community. Table 1 shows the main descriptive data regarding the individuals with ID, the practitioners and the family members who participated for every community. Data on the level of ID has been provided by the public services of Spain that are responsible for assessing persons with disability. The professionals (psychiatrists, social workers and psychologists) use the WAIS (Wechsler Adult Intelligence Scale) to measure the cognitive ability and the ICAP (in English, Inventory for service planning and individual programming) to examine adaptive behaviour and disruptive behaviour problems.

Table 1

*Descriptive statistics (percentages) of people with ID, practitioners and families*

	Andalusia	Aragon	Catalonia	Castile and Leon	Castile- La Mancha	Madrid	Galicia
<i>People with ID (n=529)</i>							
<i>Gender (%)</i>							
Male	58.60	61.90	53.80	62.90	52.60	56.50	53.30
Female	41.40	38.10	46.20	37.10	47.40	43.50	46.70
<i>Area of residence (%)</i>							
Rural	21.10	14.30	4.50	37.10	19.20	3.30	16.70
Semi-urban	34.90	85.70	34.10	-	46.20	17.40	41.70

Urban	44	-	61.40	62.90	34.60	79.30	41.70
Intellectual disability level (%)							
Borderline	10.80	14.30	5.30	2.90	19.20	4.30	1.70
Mild	31.50	23.80	36.40	62.90	33.30	30.40	20
Moderate	50.50	57.10	46.20	31.40	44.90	52.20	45
Severe and/or profound	7.20	4.80	12.10	2.90	2.60	13	33.30
Day care (%)							
Special work center	1.80	9.50	22.70	5.70	3.80	12	1.70
Occupational therapy services	76.10	81	73.50	85.70	88.50	59.80	45
Day center	8.30	9.50	3.80	2.90	3.80	17.40	43.30
Educational center	5.50	-	-	5.70	-	5.40	5
Others	8.30	-	-	-	-	5.40	1.70
Place of residence (%)							
Residence	8.70	9.50	5.30	17.60	6.60	8.70	3.40
Supervised flat	-	-	22	23.50	10.50	-	6.80
Family home	86.50	81	68.90	58.80	81.60	88	89.80
Independent home	4.80	9.50	3.80	-	1.30	3.30	-
<i>Practitioners (n=522)</i>							
Type (%)							
Direct care (day)	75	47.60	79.50	-	66.20	49.50	76.70
Direct care (night)	-	-	2.30	-	-	-	-
Direct care (physical activity)	6.70	-	-	-	13	29.70	5
Technical staff of service	13.50	42.90	17.40	100	20.80	6.60	11.70
Others	4.80	9.50	-	-	-	8.80	3.30
Educational level (%)							
Secondary education	22.1	9.50	9.10	-	17.90	6.60	16.70
University degree	58.7	42.90	64.40	94.30	51.30	42.90	41.70
Higher university degree	1.90	-	11.40	5.70	14.10	5.50	21.70
Others	17.30	47.60	15.20	-	16.70	45.10	20
<i>Family (n=462)</i>							
Relation with person with ID (%)							
Parent	72.40	42.90	66.40	54.50	81.20	83.10	74.60
Sibling	21.80	52.40	21.80	36.40	15.90	12	22
Other family member	4.60	4.80	2.70	-	1.40	4.80	3.40
Legal tutor	1.10	-	9.10	9.10	1.40	-	-
Educational level (%)							
No studies	19.80	4.80	6.40	-	20	12.20	6.80
Primary education	41.90	23.80	42.20	60	47.10	20.70	52.50

Secondary education	18.60	38.10	26.60	30	15.70	24.40	18.60
University studies	16.30	14.30	18.30	10	11.40	32.90	15.30
Others	3.50	19	6.40	-	5.70	9.80	6.80
Place of residence (%)							
Rural	19.50	14.30	14.50	23.30	21.40	3.60	16.90
Semi-urban	43.70	85.70	36.40	3.30	42.90	15.70	45.80
Urban	36.80	-	49.10	73.30	35.70	80.70	37.30

## Instrument

The Spanish version of the POS (Carbó-Carreté *et al.* 2015) aims to assess QoL in people with ID on the basis of three factors and eight domains validated in a series of cross-cultural studies: (1) Independence, composed of Personal Development and Self-determination; (2) Social Participation, which includes Interpersonal Relations, Social Inclusion, and Rights; and (3) Well-being, which encompasses Emotional Well-being, Physical Well-being, and Material Well-being (Jenaro *et al.* 2005; Schalock *et al.* 2005; Wang *et al.* 2010).

This scale is divided, as mentioned above, into three information sources: (a) self-reporting, in which the individual answers on his/her own; therefore, this source assesses the subjective perspective of QoL; (b) reporting by the practitioner, which assesses the individual's experiences and circumstances from the viewpoint of direct care staff or a service technician; and (c) family reporting, in which the indicators are scored from a family member's perspective. As mentioned above, the self-report and the report-of-others have the same contents but are expressed differently, according to the person being interviewed. For example, in the self-report the items are expressed as follows: "Do people take your decisions seriously?", "What about your health in general? How do you feel?". In report-of-others, instead, they are expressed as follows: "To what degree are the decisions made by the person respected (regardless of the specific decision)?", "In general, how would you evaluate the physical health of this person?" Every dimension has 6 items; therefore, a total of 48 items are answered for the scale as a whole. Every item is assessed using a 3-point Likert scale. Scores are obtained through interviews conducted by an interviewer with previous training regarding the theoretical model and proper administration of the scale. Outcomes are obtained for every dimension and for the three factors. For every dimension, the sum of all of the scores from the 6 items is obtained using the following calculation: (3) =

*always*; (2) = *sometimes*; and (1) = *rarely or never*. After the dimensions of every factor are totalled, a final score is calculated for each factor. The self-report scale takes approximately one hour to administer, whereas the report by the practitioner and the family take about forty-five minutes each.

The reliability study provides appropriate values for the first and second order factors with  $\alpha$  values greater than .82. The construct validity analysis provides an adjustment of the theoretical model, particularly regarding the assessments from practitioners (Table 2). Pearson's correlations between factors are also coherent. For the first-order factor the lowest values were between rights and social inclusion domains ( $r = .32, p < .001$ ), and the highest between interpersonal relations and personal development ( $r = .73, p < .001$ ), both correlations in the self-report answered by the individual with ID. In the second-order factor the lowest values were between well-being and social participation ( $r = .42, p < .001$ ) for the self-report and the highest between social participation and independence ( $r = .77, p < .001$ ) for the family report (Carbó-Carreté *et al.* 2015).

## **Procedure**

Organizations that provide services were asked to participate by the Spanish Confederation of Organizations for the People with Intellectual Disability (*Plena Inclusión*) and by logistic support from the delegations in every autonomous community. The organizations that agreed to participate provided day services (special work centres, occupational therapy, day centres), and most of them also provided residential services (supervised flats, residences) for adults with ID.

In every autonomous community, specific training was provided regarding the application of the POS to those practitioners who would participate as interviewers. The training was about one or two days long depending on the number of assistants and it was made by the authors who were working on the translation and validation of the Spanish version. These trainers knew the QoL model well and had previously knowledge and had reviewed the pilot test of the Spanish POS along with the authors of the original version. By doing so, we could guarantee that the instrument's application would be in concordance with the original authors' guidelines. At the end of the training sessions, the interviewers were provided all of the materials (scale forms and item descriptors) that were needed to administer the scale in every centre. Once the person



with ID was identified, the practitioner and the family were contacted and defined as well. For persons with ID the requirements were that the participants were over 16 years old and that they were assisted by one of the daily services. To guarantee the deep knowledge of the practitioner, the inclusion criterion was that the interviewee had to have assisted the person with ID for at least three months and have observed him or her in their natural settings. Along with the questionnaires, to follow the guidelines of the Ethical Committee, the interviewers received an informed consent form so that it could be read and signed by all of the participants.

The trained practitioners interviewed a total of 670 participants, 529 of whom obtained the participation of the three sources of information. The reason to exclude some participants were mainly that they did not answer enough dimensions to examine the psychometric properties of the scale.

### **Statistical Analysis**

In this work, all of the items of every domain were operationally defined using a graduated Likert scale. That is, the metric base of each item was clearly ordinal. If we assume that each dimension of the proposed scale constitutes in itself a specific factor, we can conceive of each dimension as a one-dimensional structure and, therefore, analyse them in a specific manner individually.

The IRT allows to estimate the extent to exist a correct discrimination of each of the items subjected to study. In this way, it is possible to evaluate whether each of the evaluation sources present agreement or disagreement, not in the value obtained in each item (that would be solved by simple comparison techniques), but in the overall response pattern that each source presents in each item. So, we get two kinds of important information. The first is to know if each item has discrimination capacity according to each source of evaluation and the second, to analyze the patterns of each source to identify both similarities and dissimilarities. Both provide relevant information for the evaluator. The adequacy is to know the discrimination capacity of each item in each source (IRT curves) and its significance (parameters); while allowing us to identify items of similar pattern between sources and others of non-similar pattern.

From the perspective of the IRT, there are several possibilities of models applicable to one-dimensional structures with ordinal and graduated metrics. Among them, the model of Samejima (1969) is likely among the best known, and it was used with some later modifications, such as those proposed by Muraki (1990) for the Likert

scale. The Samejima model proposes that the probability for a given response category in a given item can be estimated using this logistic function:

$$P_{jk}(\theta) = \frac{1}{1 + \exp[Da_j (\theta - b_{jk})]} - \frac{1}{1 + \exp[Da_j (\theta - b_{jk+1})]}$$

where  $\theta$  is the value of each response category,  $a_j$  is the slope estimate for each item,  $b_{jk}$  is the cut-off point between categories, and  $D$  is a constant equivalent to 1.7. The aforementioned modification of Muraki (1990) was applied in cases in which the number of categories was constant among the items, as in our work. The parameter  $b_{jk}$  was decomposed into two parameters:  $b_j$  indicating the item location parameter; and  $c_k$ , which is the location parameter of each cut-off point between categories, which are constant among the items. This procedure, as Hernández *et al.* (2000) noted, allowed us to order the items relative to the measured psychological dimension.

Different approaches have been proposed for the estimation of the previous logistic function, and in our case, we chose to estimate the parameter  $b_j$  from the logic of the Structural Equation Models (SEMs). To this end, we used the determination of a Confirmatory Factor Analysis (CFA) model, according to the general expression, which is widely known:

$$x = \Lambda_x \cdot \zeta + \delta.$$

where  $X$  is the vector of response values,  $\Lambda_x$  is the matrix of factor loads, all of which are different from 0,  $\zeta$  is the latent factor vector, and finally,  $\delta$  is the measurement error. The estimates applied to the categorical variables have been described by Baker & Kim (2004) and by du Toit (2003). These values of  $\Lambda_x$  constitute TRI estimates of the parameter  $b_j$  described above; therefore, they assume an estimation of the location of the item in the continuum defined by the one-dimensional scale. Thus, we analysed, according to this assumption, each of the dimensions of the scale (Personal Development, Self-determination, Interpersonal Relations, Social Inclusion, Rights, Emotional Well-being, Physical Well-being, Material Well-being) and each source of information (self-reporting, reporting by practitioners, reporting by families), using the

modified CFA model as described above with Maximum Likelihood (ML) estimations adapted to categorical data, according to the Muthén & Muthén (1998-2007) options.

## Results

According to the statistical procedures mentioned above, we estimated the adjustment values of all of the CFA models obtained for each dimension and source of information. Thus, 24 different models were obtained that should fit the matrix of ordinal correlations between items. Table 2 shows the  $\chi^2$  fit values of the ML estimates for each of these 24 models. We excluded the remainder of the usual adjustment indicators since, with only the estimation of  $\chi^2$ , the one-dimensional adjustment of each model was already evident.

Table 2

*Adjustment data of  $\chi^2$  for each CFA model, including the degrees of significance*

	Self-report	Report by practitioner	Report by family
Personal Development	489.239 $p = 1$	401.675 $p = 1$	524.992 $p = 1$
Self-determination	394.189 $p = 1$	282.676 $p = 1$	319.913 $p = 1$
Interpersonal relations	285.588 $p = 1$	372.932 $p = 1$	262.264 $p = 1$
Social inclusion	581.977 $p = .9998$	495.119 $p = 1$	524.360 $p = 1$
Rights	499.622 $p = 1$	540.314 $p = 1$	478.885 $p = 1$
Emotional well-being	269.887 $p = 1$	191.645 $p = 1$	209.492 $p = 1$
Physical well-being	350.126 $p = 1$	378.507 $p = 1$	372.525 $p = 1$
Material well-being	573.066 $p = .9999$	585.474 $p = .9997$	436.600 $p = 1$

The adjustment results, as mentioned, left no room for doubt about the adjustment to the one-dimensional model of each factor and source of information. Thus, we estimated the parameters of  $\mathcal{A}_x$  to obtain the parameter  $b_j$  for each model, using a modified Samejima model for graded Likert scales. Table 3 shows the estimates for each of the parameters with their corresponding statistical significance.

Table 3

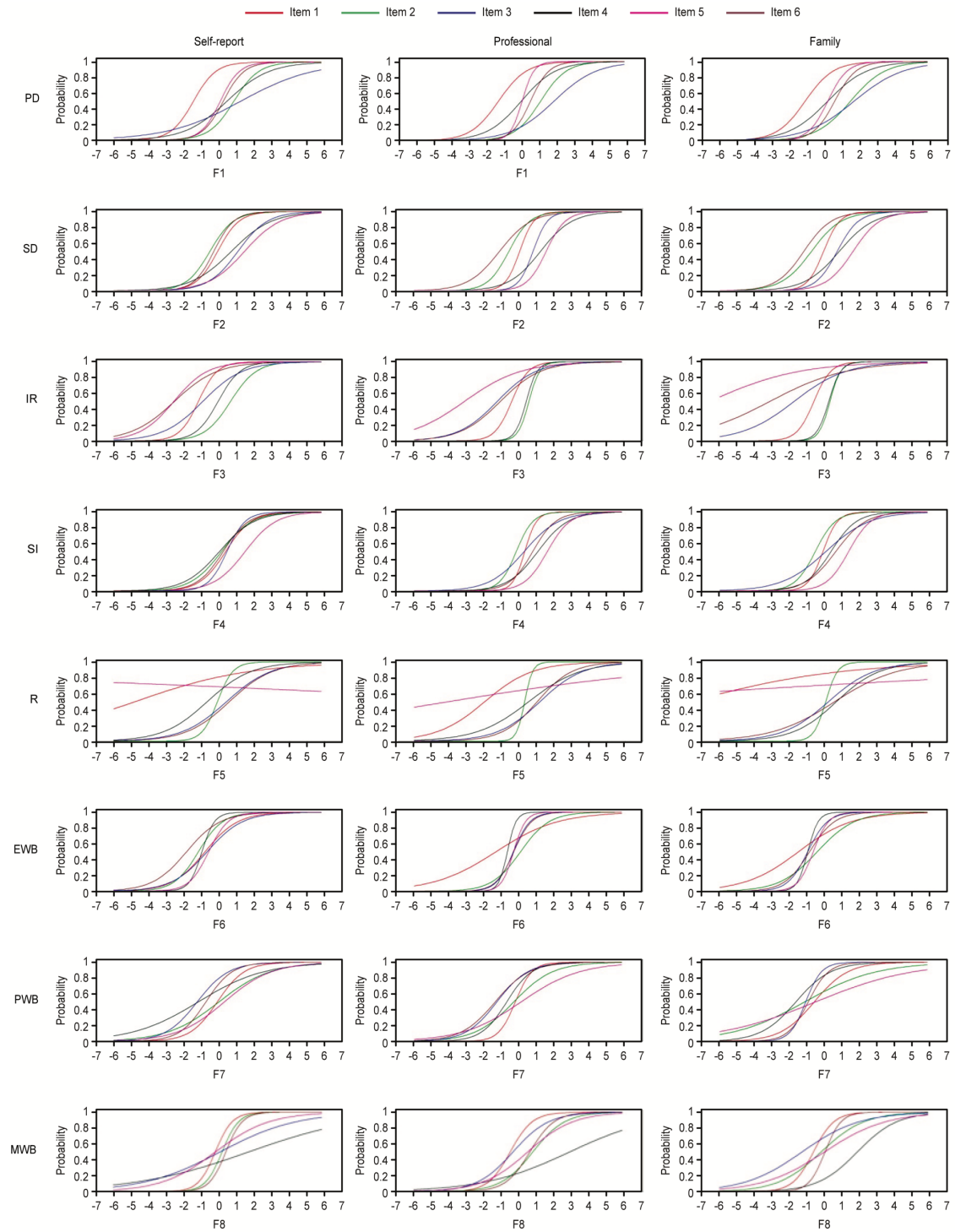
*Parameter estimation of Samejima's model applied to each information source*

	Self-report	Report by practitioner	Report by family
Personal Development (PD)			
PD1	1.572 (0.298)	1.257 (0.208)	1.252 (0.255)
PD2	1.417 (0.210)	1.363 (0.174)	1.024 (0.177)
PD3	0.483 (0.117)	0.829 (0.127)	0.715 (0.143)
PD4	0.836 (0.140)	1.024 (0.147)	0.898 (0.164)
PD5	1.841 (0.269)	2.976 (0.481)	1.815 (0.316)
PD6	1.695 (0.240)	1.907 (0.223)	1.648 (0.267)
Self-Determination (SD)			
SD1	1.640 (0.228)	2.489 (0.313)	2.183 (0.341)
SD2	1.498 (0.224)	1.675 (0.211)	1.179 (0.192)
SD3	1.191 (0.165)	2.448 (0.303)	1.659 (0.225)
SD4	0.873 (0.149)	1.113 (0.163)	1.023 (0.170)
SD5	0.975 (0.144)	1.842 (0.209)	1.315 (0.185)
SD6	1.760 (0.262)	1.111 (0.169)	1.281 (0.217)
Interpersonal Relations (IR)			
IR1	1.871 (0.346)	2.095 (0.247)	2.072 (0.288)
IR2	1.363 (0.210)	3.241 (0.457)	3.259 (0.536)
IR3	0.885 (0.175)	0.901 (0.147)	0.675 (0.149)
IR4	1.594 (0.259)	3.196 (0.455)	2.947 (0.441)
IR5	1.043 (0.267)	0.621 (0.165)	0.390 (0.217)
IR6	0.813 (0.213)	0.871 (0.148)	0.461 (0.148)
Social Inclusion (SI)			
SI1	1.366 (0.209)	2.971 (0.437)	2.413 (0.385)
SI2	1.148 (0.180)	2.116 (0.257)	1.736 (0.259)
SI3	1.916 (0.295)	1.012 (0.140)	0.801 (0.139)
SI4	1.011 (0.154)	1.167 (0.142)	1.371 (0.186)
SI5	1.141 (0.162)	1.547 (0.182)	1.586 (0.218)
SI6	1.361 (0.203)	1.527 (0.182)	1.177 (0.167)

Rights (R)			
R1	0.307 (0.164)	0.736 (0.165)	0.231 (0.185)*
R2	2.287 (0.722)	4.398 (2.151)	2.988 (1.605)
R3	0.788 (0.164)	0.823 (0.137)	0.771 (0.197)
R4	0.802 (0.170)	0.679 (0.127)	0.816 (0.187)
R5	-0.045 (0.135) *	0.144 (0.114)*	0.061 (0.142)*
R6	0.818 (0.166)	1.008 (0.145)	0.563 (0.140)
Emotional Well-being (EW)			
EW1	1.073 (0.176)	0.566 (0.126)	0.660 (0.154)
EW2	1.455 (0.235)	1.225 (0.163)	0.941 (0.161)
EW3	0.992 (0.168)	2.093 (0.251)	1.783 (0.277)
EW4	2.819 (0.554)	4.093 (0.767)	3.392 (0.714)
EW5	1.818 (0.268)	2.980 (0.416)	2.314 (0.354)
EW6	1.041 (0.194)	1.973 (0.241)	1.407 (0.222)
Physical Well-being (PW)			
PW1	1.299 (0.260)	2.123 (0.412)	1.043 (0.218)
PW2	0.685 (0.151)	0.906 (0.153)	0.489 (0.158)
PW3	1.121 (0.233)	1.135 (0.213)	2.243 (0.521)
PW4	0.536 (0.152)	1.394 (0.223)	1.008 (0.228)
PW5	0.746 (0.160)	0.600 (0.133)	0.359 (0.141)
PW6	1.336 (0.262)	1.035 (0.192)	1.734 (0.325)
Material Well-being (MW)			
MW1	2.084 (0.284)	1.592 (0.273)	1.715 (0.294)
MW2	2.118 (0.285)	1.102 (0.176)	0.820 (0.160)
MW3	0.461 (0.116)	0.981 (0.165)	0.614 (0.147)
MW4	0.310 (0.115)	0.410 (0.132)	0.890 (0.204)
MW5	0.654 (0.126)	0.778 (0.146)	0.574 (0.140)
MW6	2.148 (0.286)	1.448 (0.217)	2.193 (0.436)

Note: In brackets the standard error values of each estimate. \* Indicates non-statistically significant parameter.

Based on the values obtained, it is clear that in all of the cases (dimensions and sources) a location parameter ( $b_j$ ) is obtained, which indicates that each item has a different position, and its behaviour in the estimation of the total scale must be considered. All of these values are different from zero except for the four parameters that have non-significant results found on item 1 for Rights in the family information source (0.231) and in item 5 on the same dimension for the three sources of information (-0.045; 0.144 and 0.061). Moreover, a very simple inspection of the parameter's values shows differences between them in some sources and factors. It can be interpreted as an item with high discrimination capacity according with the function of each source. Therefore, the discrimination function of each item is not the same among the three sources. Thus, it is important to determine the exact locations of the items of each factor and the source of information to determine the psychological distances of the latent factors. For this purpose, we show the cumulative characteristic curves for the three response categories of the items for each factor and source of information. In every curve, the cumulative distribution of each item is presented in the three categories defined as response criteria: (3) = *always*; (2) = *sometimes*; and (1) = *rarely or never*.



*Figure 1.* Characteristics curves of all items for each latent factor and information source.

## Discussion

Figure 1 shows the behaviours of the characteristic curves of the items (CCI) on each dimension and according to each source of information involved in the assessment of



the QoL of the person with ID. It is observed that, in some dimensions, such as Personal Development, Social Inclusion and Rights, the assessments performed by the person with ID are similar to those reported by the others. However, other dimensions presented clear differences between the values of the person with ID and those of the practitioners and families, as in Self-determination and Interpersonal Relations. The Well-being factor, which includes Emotional, Physical and Material Well-being, should be considered carefully because it showed different operating items among the three groups of informants. Therefore, given the relevance of each dimension in the assessment of personal results (Carbó-Carreté *et al.* 2015), it is necessary to dwell on each of the dimensions and to observe the behaviours of different items.

In the Personal development dimension, the behaviour of the items was similar among the three informants. In all three cases, item 1 showed the least difficulties, likely because most of the sample consisted of autonomous people or those with few needs for support in daily living activities. In contrast, items 5 and 6 indicated higher discriminating capacity according to their curves. The content of both items was very clear and easy to interpret (access to information and use of computers, cell phones, etc.), which was not the case with the other items, which were more difficult to answer, and the options associated with the QoL, thus showing a low level of discrimination. As shown in Figure 1 for practitioners and families, items 2 and 3 were less discriminating; however, in people with ID items 3 and 4 were less discriminating. Among the three informants, it was evident that item 3 had the greatest difficulty and was less discriminative. The content of this item (“Are you learning skills to do more things, or do you follow courses/are you engaged in some type of education?”) tended to lend ambiguity, given the difficulty in understanding types of skills. Often, practitioners and families and even people with the same ID seemed to respond with regard to whether they were involved in some activity programme organised by the service without considering the existence of an individual plan to develop a particular ability.

Regarding the dimension of Self-determination, the items assessed by people with ID showed different behaviours than those answered by practitioners and families. In the first group, items 1, 2 and 6 showed strong consistency among them, as did, to a lesser degree, items 3, 4 and 5. This finding was not surprising because it considered the content of the items. Items 1, 2 and 6, although slightly different, all involved the possibility that the same person must choose and express his or her opinion. Although versions of practitioners and families did not show consistency among items, they

agreed with the people with ID about the behaviour of the items. For the three informants, items 1, 2 and 6 were less difficult to answer, and item 5 was the most problematic. This item valued control of their own money and led to confusion due to the many issues that arose (money allocated for weekly costs, money they earn each month, money they receive from public administration, among others).

Regarding the Interpersonal Relations dimension, items 3, 5, and 6 deserve special attention because they had low discriminatory capacity, and even considering the curves of the families' data, it could be necessary to conduct an exhaustive review of all of these items. Actually, item 5 attracted attention ("Does the family treat the person with unconditional respect and dignity or express his/her importance in other ways?") because it was the item that functioned worst for families and practitioners, and it worked slightly better in people with ID. Most of the sample lived with their families and likely had trouble understanding this item

In Social Inclusion, item 5 was the item that clearly showed the most difficulty in providing the expected response. This item assessed a complex concept ("Do people from the community do things for you including visiting you and taking you places?") and required some skill to interpret and answer correctly. In relation to the other items, in people with ID, the items had very high consistency among them, and the ability to discriminate was acceptable in general. This finding was not the case in the group of practitioners and families, and there was more dispersion among items, with items 1 and 2 the most discriminative.

The results of the Rights dimension indicated that all of the items required careful analysis since the functioning of the items was unsatisfactory. Among the three informants, item 5 did not fit the expected pattern, nor did item 1, mainly in people with ID and in the family group. Among the three informants, item 2 was the most discriminative ("Do you control the key to your home or apartment?"), likely because it was a clear indicator of rights that must be enjoyed by anyone regardless of their level of ID.

Regarding Emotional Well-being, the only coincidence that occurred among the three groups was the high capacity for discriminating shown in item 4 ("Would you say you are a happy person?"). However, other items behaved in a similar manner for practitioners and families, although they differed for people with ID. The data obtained using reporting by others (practitioners and families) presented a certain consistency among the items and poor discriminatory capacity, except for the first two items. In

these two groups, the first item (“How would you rate the safety and security of the person’s daily environments?”) functioned incorrectly, probably because it tended to evaluate only safety in physical environment and did not consider third persons or unfavourable environments. However, in people with ID, most of the items functioned correctly, and item 6 (“Do you have basic trust in people who are important to you?”) showed a higher probability of responding correctly.

In the Physical Well-being dimension, items 2 and 5 should be reviewed due to their low functioning in the family group, as well as item 4, because it had low capacity to discriminate in the people with ID. Basically, in this dimension, the diversity of the behaviours of items amongst the three informants was emphasised. In the practitioner group, item 1 was the most discriminative, in families, item 3 was the most discriminative, and in people with ID, items 1, 3 and 6 worked in a similar fashion.

Finally, Material Well-being showed the poorest behaviour in item 4 (“Do you have a paid job?”), specifically in people with ID and practitioners. Items 3 and 5 showed little discrimination in people with ID and their families. The items that were more discriminative according to the three groups of informants were 1 and 6, also including item 2, but only in people with ID.

After reviewing each of the dimensions, we conclude this project by emphasizing three fundamental aspects. First, as had been noted previously by referenced authors in this field (Balboni *et al.* 2013, Claes *et al.* 2012a; Schmidt *et al.* 2010; Simões & Santos 2016), in this paper, we emphasised the importance of self-reporting and reporting by others on the same scale; in others words, it is important to conduct evaluations among the three groups of informants using the same items. Based on these results we support the fact that each interviewee contributes essential information and no interviewee is replaceable (Claes *et al.* 2012a). We also confirm that, in case the person with ID cannot answer the self-report part by him- or herself, it may not be reported by other informants (van Loon, 2015). Focusing on the self-report part, we have the conditions to examine if items are adequate for all people with ID independently of their severity. This question provides us with a new issue to study, and possibly not only for the self-report but also for the report-of-others. Second, when interpreting the evaluations obtained in each dimension, it is necessary to determine the similarities and differences that occur between the informants and to examine in depth the differences in each particular case. Finally, the third point maintains a close relationship with the POS used. Due to the low functioning of certain items, there is a

clear need to review the content of these items, especially those that did not work in any of the three groups of informants. At least, we should be cautious in assessments resulting from the use of the POS without some considerations regarding differences between item distributions according to the information source.

## References

- Baker F.B. & Kim S. (2004) Item response theory. Parameter estimation techniques. Second edition. Marcel Dekker, Inc., New York.
- Balboni G., Coscarelli A., Giunti G. & Schalock R.L. (2013) The assessment of the quality of life of adults with intellectual disability: The use of self-report and report of others assessment strategies. *Research in Developmental Disabilities* **34**, 4248-4254. doi: 10.1016/j.ridd.2013.09.009.
- Bonham G.S., Basehart S., Schalock R.L., Marchand C.B., Kirchner N. & Rumenap J. M. (2004) Consumer-based quality of life assessment: The Maryland Ask Me! Project. *Mental Retardation* **42**, 338– 355. doi: 10.1352/0047-6765(2004)42<338:CQOLAT>2.0.CO;2
- Brown I., Hatton C. & Emerson E. (2013) Quality of life indicators for individuals with intellectual disabilities: extending current practice. *Intellectual and Developmental Disabilities* **51**, 316-332. doi: 10.1352/1934-9556-51.5.316
- Carbó-Carreté M., Guàrdia-Olmos J. & Giné C. (2015) Psychometric properties of the Spanish version of the Personal Outcomes Scale. *International Journal of Clinical and Health Psychology* **15**, 236-252. doi: 10.1016/j.ijchp.2015.04.002.
- Claes C., Vandevelde S., Van Hove G., van Loon J., Verschelden G. & Schalock R.L. (2012a) Relationship between self-report and proxy ratings on assessed personal quality of life-related outcomes. *Journal of Policy and Practice in Intellectual Disabilities* **9**, 159-165. doi: 10.1111/j.1741-1130.2012.00353.x.
- Claes C., Van Hove G., Vandevelde S., van Loon J. & Schalock R.L. (2012b) The influence of support strategies, environmental factors, and client characteristics on quality of life-related outcomes. *Research in Developmental Disabilities* **33**, 96-103. doi:10.1016/j.ridd.2011.08.024
- Cummins R.A. (2005) Caregivers as managers of subjective wellbeing: A homeostatic perspective. *Journal of Applied Research in Intellectual Disabilities* **18**, 335– 344. doi: 10.1111/j.1468-3148.2005.00267.x

- du Toit M. (2003) IRT from SSI: Bilog-mg multilog parscale testfact. Scientific Software International, Inc., Lincolnwood, IL
- Finlay W.M. & Lyons E. (2001). Methodological issues in interviewing and using self-report questionnaires with people with mental retardation. *Psychological assessment* **13**, 319-335. doi: 10.1037/1040-3590.13.3.319
- Finlay W.M. & Lyons E. (2002). Acquiescence in interviews with people who have mental retardation. *Mental retardation* **40**, 14-29. doi 10.1352/00476765(2002)040<0014:AIWPW>2.0.CO;2
- Hartley S.L. & MacLean J.W.E. (2006) A review of the reliability and validity of Likert-type scales for people with intellectual disability. *Journal of Intellectual Disability Research* **50**, 813–827. doi: 10.1111/j.1365-2788.2006.00844.x
- Hernández A., Muñiz J. & García E. (2000) Comportamiento del modelo de respuesta graduada en función del número de categorías de la escala. *Psicothema* **12**, (S. 2).
- Janssen C.G.C., Schuengel C. & Stolk J. (2005) Perspectives on quality of life of people with intellectual disabilities: the interpretation of discrepancies between clients and caregivers. *Quality of Life Research* **14**, 57-69. doi: 10.1007/s11136-004-1692-z.
- Jenaro C., Verdugo M.A., Caballo C., Balboni G., Lachapelle Y., Otrebski, W. et al. (2005) Cross-cultural study of person-centred quality of life domains and indicators: A replication. *Journal of Intellectual Disability Research* **49**, 734-739.
- Li C., Tsoi E. W., Zhang A. L., Chen S. & Wang, C. J. (2013) Psychometric properties of self-reported quality of life measures for people with intellectual disabilities: A systematic review. *Journal of Developmental and Physical Disabilities* **25**, 253-270. doi: 10.1007/s10882-012-9297-x
- Muraki E. (1990) Fitting a polytomous item response model to Likert-type data. *Applied Psychological Measurement* **14**, 59-71.
- Muraki E. & Bock R.D. (1998) PARSCALE. IRT item analysis and test scoring for rating-scale data. Scientific Software International, Chicago.
- Muthén L.K. & Muthén B.O. (1998-2007) Mplus User's Guide. Fifth Edition. Muthén & Muthén, Los Angeles, CA.
- Perry, J., & Felce D. (2002) Subjective and objective quality of life assessment: Responsiveness, response bias, and resident: Proxy concordance. *Mental*

- Retardation* **40**, 445–456. doi:  
10.1352/00476765(2002)040<0445:SAOQOL>2.0.CO;2
- Perry J. & Felce D. (2005) Factors associated with outcomes in community group homes. *American Journal on Mental retardation* **110**, 121-135. doi:  
10.1352/0895-8017(2005)1102.0.CO;2.
- Reinders H.S. & Schalock R.L. (2014) How organizations can enhance the quality of life of their clients and assess their results: the concept of QOL enhancement. *American Journal on Intellectual and Developmental Disabilities* **119**, 291-302. doi: 10.1352/1944-7558-119.4.291
- Samejima F. (1969) Estimation of latent ability using a response pattern of graded scores. *Psychometrika Monograph* **17**.
- Schalock R.L., Bonham G.S. & Verdugo M.A. (2008a) The conceptualization and measurement of quality of life: Implications for program planning and evaluation in the field of intellectual disabilities. *Evaluation and Program Planning* **31**, 181-190. doi: 10.1016/j.evalprogplan.2008.02.001.
- Schalock R.L., Brown I., Brown R., Cummins R.A., Felce D., Matikka L. et al. (2002) Conceptualization, measurement, and application of quality of life for persons with intellectual disabilities: Results of an international panel of experts. *Mental Retardation* **40**, 457–470. doi:  
10.1352/00476765(2002)040<0457:CMAAOQ>2.0.CO;2
- Schalock R.L. & Felce D. (2004) Quality of life and subjective well-being: Conceptual and measurement issues. In: *International handbook of applied research in intellectual disabilities* (eds E. Emerson, C. Hatton, T. Thompson & T.R. Parmenter), pp. 261-279. Wiley, London.
- Schalock R.L., Gardner J.F. & Bradley V.J. (2007) Quality of life for persons with intellectual and other developmental disabilities: Applications across individuals, organizations, communities and systems. American Association on Intellectual and Developmental Disabilities, Washington, DC.
- Schalock R.L. & Verdugo M.A. (2002) Handbook on quality of life for human service practitioners. American Association on Mental Retardation, Washington, DC.
- Schalock R.L. & Verdugo M.A. (2012) *A leadership guide for today's disabilities organizations: Overcoming challenges and making change happen*. Baltimore: Brookes Publishing Co.

- Schalock R.L., Verdugo M.A., Bonham G.S., Fantova F. & Van Loon J. (2008b) Enhancing personal outcomes; organizational strategies, guidelines, and examples. *Journal of Policy and Practice in Intellectual Disabilities* **5**, 276-285. doi: 10.1111/j.1741-1130.2007.00135.x.
- Schalock R.L., Verdugo M.A., Jenaro C., Wang M., Wehmeyer M.L., Jiancheng X. et al. (2005) Cross cultural study of quality of life indicators. *American Journal of Mental Retardation* **110**, 298-311.
- Schmidt S., Power M., Green A., Lucas-Carrasco R., Eser E., Dragomirecka E. et al. (2010) Self and proxy rating of quality of life in adults with intellectual disabilities: Results from the DISQOL study. *Research in developmental disabilities* **31**, 1015-1026. doi:10.1016/j.ridd.2010.04.013.
- Schwartz C. & Rabinovitz S. (2003) Life satisfaction of people with intellectual disability living in community residences: Perceptions of the residents, their parents and staff members. *Journal of Intellectual Disability Research* **47**, 75–84. doi: 10.1046/j.1365-2788.2003.00436.x
- Simões C. & Santos S. (2016) The quality of life perceptions of people with intellectual disability and their proxies. *Journal of Intellectual & Developmental Disability* **41**, 311-323.
- Stancliffe R.J. (2000) Proxy respondents and quality of life. *Evaluation and Program Planning* **23**, 89–93. doi: 10.1016/S0149-7189(99)00042-7
- Thompson J.R., Bradley V., Buntinx W.H.E., Schalock R.L., Shogren K.A., Snell M.E. et al. (2009) Conceptualizing supports and the support needs of people with intellectual disability. *Intellectual and Developmental Disabilities* **47**, 135–146. doi: 10.1352/1934-9556-47.2.135.
- van Loon J.H. (2015) Planes individuales de apoyo: mejora de los resultados personales. *Siglo Cero* **46**, 25-40. doi:http://dx.doi.org/10.14201/scero20154612540.
- van Loon J., Van Hove G., Schalock R.L. & Claes C. (2008) *Personal Outcomes Scale: A Scale to Assess an Individual's Quality of Life*. Middelburg: Stichting Arduin and Gent: University of Gent.
- Verdugo M.A., Gómez L.E., Arias B., Santamaría M., Clavero D. & Tamarit J. (2013) Evaluación de la calidad de vida en personas con discapacidad intelectual: la Escala INICO-FEAPS [Assessment of quality of life in persons with intellectual disability: INICO-FEAPS Scale]. *Siglo Cero* **44**, 6–20.

- Verdugo M. A., Schalock R. L., Keith K. D. & Stancliffe, R. J. (2005) Quality of life and its measurement: Important principles and guidelines. *Journal of Intellectual Disability Research* **49**, 707–717. doi:10.1111/j.1365-2788.2005.00739.x
- Wang M., Schalock R.L., Verdugo M.A. & Jenaro C. (2010) Examining the factor structure and hierarchical nature of the quality of life construct. *American Journal of Intellectual and Developmental Disabilities* **115**, 218–233. doi:10.1352/1944-7558-115.3.218.